Dietary Patterns Predict Mortality in a National Cohort: The National Health Interview Surveys, 1987 and 1992[1]

Ashima K. Kant,[2] Barry I. Graubard,* and Arthur Schatzkin†

*Department of Family, Nutrition, and Exercise Sciences, Queens College of the City University of New York, Flushing, NY; *The Division of Cancer Epidemiology and Genetics, Biostatistics Branch, National Cancer Institute, National Institutes of Health, Bethesda, MD; and †The Division of Cancer Epidemiology and Genetics, Nutritional Epidemiology Branch, National Cancer Institute, National Institutes of Health, Bethesda, MD*

ABSTRACT We examined the association of mortality and dietary patterns using data from the National Health Interview Surveys of 1987 and 1992 (n = 10,084), aged ≥45 y at baseline (with 2287 deaths due to all causes over 5.9 median years of follow-up). The ~60-item FFQ administered at baseline was examined for mentions of foods and dietary behaviors recommended in current dietary guidance (fruits, vegetables, lean poultry and alternates, low-fat dairy, and whole grains), and the resulting patterns were expressed as follows: *1*) a Recommended Foods and Behavior Score (RFBS), *2*) factor scores from factor analysis, and *3*) clusters from cluster analysis. The multivariate-adjusted relative risk (RR) of mortality for each of the 3 types of dietary patterns was examined using Cox proportional hazards regression analysis. In men, RR of all-cause mortality was 0.72 (95% CI: 0.56, 0.92, P for trend < 0.001) for RFBS, and 0.74 (95% CI: 0.57, 0.95, P for trend = 0.002) for the fruit-vegetable-whole grain factor score when comparing extreme quartiles. Membership in 1 of the 4 clusters also was associated with lower risk in men (RR = 0.82, 95% CI: 0.66, 1.01). For women, the RFBS was a modest inverse predictor of mortality after multivariate adjustment (RR = 0.80, 95% CI: 0.61, 1.04, P for trend = 0.04), but estimates for factor and cluster patterns were attenuated. The population-attributable fraction due to diet was 0.16 in men and 0.09 in women. Dietary patterns characterized by compliance with prevailing food-based dietary guidance were associated with a lower risk of all-cause mortality. J. Nutr. 134: 1793–1799, 2004.

KEY WORDS: • *dietary patterns* • *diet quality* • *factor analysis* • *cluster analysis* • *NHIS* • *all-cause mortality*

Present day dietary guidance reflects our understanding of the roles of various dietary constituents in altering the risk of chronic diseases (1–3). Evidence for these diet and health associations has traditionally come from studies that focused on single nutrients, foods, or food groups in relation to particular health outcomes (1). In recent years, dietary patterns, loosely defined as multiple dietary attributes operationalized as a single exposure, have emerged as an alternative or adjunct to traditional approaches for the study of the association of diet and health (4–6). The study of diet and disease associations with the dietary patterns approach has intuitive appeal because it tends to reflect the multidimensional nature of dietary behaviors and the intercorrelation of nutrient metabolism and utilization in vivo (4,5). Consequently, a number of papers that examined health outcome in relation to dietary patterns derived from 1 of 3 possible approaches have appeared in the literature (4,5). The prevailing approaches for determining dietary patterns include simple or complex scores derived from assessment of desirable dietary characteristics and statistical techniques of factor and cluster analysis (5).

Recently, we reported that dietary patterns characterized by at least weekly intake of foods recommended in current dietary guidance, assessed by a Recommended Foods Score (RFS),[3] were associated with a lower risk of all-cause mortality in a large screening cohort of women (Breast Cancer Detection and Demonstration Project or BCDDP) (7). However, given the nonrepresentative nature of the BCDDP screening cohort, we wished to extend our previous work and examine the following: *1*) whether dietary patterns consistent with current dietary guidance predict the risk of mortality in a nationally representative cohort, and *2*) whether dietary patterns derived from different approaches are comparable in predicting the risk of mortality.

SUBJECTS AND METHODS

We used data from the National Health Interview Surveys (NHIS) 1987 and 1992 for the present study. The NHIS is conducted

[1] Supported in part by National Cancer Institute grant CA91316.
[2] To whom correspondence should be addressed. E-mail: akant@qc.edu.

[3] Abbreviations used: BCDDP, Breast Cancer Detection and Demonstration Project; NCHS, National Center for Health Statistics; NDI, National Death Index; NHIS, National Health Interview Surveys; PCA, principal components analysis; RFBS, Recommended Foods and Behavior Score; RFS, recommended foods score.

Manuscript received 1 December 2003. Initial review completed 27 January 2004. Revision accepted 5 April 2004.

annually by the National Center for Health Statistics (NCHS) to provide information on health and disability within the civilian noninstitutionalized population of the United States (8). The NHIS is a cross-sectional household survey with continuous sampling and in-person interviewing throughout the year. The NHIS sample is a stratified multistage area probability sample of U.S. households (8).

The NHIS fielded from 1982 to 1996 included 2 parts: *1)* a set of basic health and demographic items (core questionnaire), and *2)* 1 or more sets of supplemental questions on topical health issues (8). The NHIS core questionnaire was unchanged through this period, whereas the current health topics in the supplement changed on the basis of data needs (8). In the 1987 and 1992 survey years, the supplement part of the NHIS questionnaire comprised a set of questions on cancer risk factors along with an abbreviated Block/NCI FFQ.

For the NHIS 1987 and 1992, the interviewed sample for the core questionnaire included 122,859 and 128,412 persons, respectively (9,10). The response rate for the core questionnaire was 95% in 1987 and 96% in 1992 (9,10). For the Cancer Risk Factor supplement fielded with NHIS 1987, a total of 22,080 questionnaires were completed, representing 1 adult person > 18 y of age, selected randomly from each household. For the Cancer Risk Factor supplement fielded with NHIS 1992, a total of 12,005 questionnaires were completed between January and July 1992. The supplement was not fielded after the first half of 1992 due to budgetary constraints. The response rate for the identified eligible respondents for the supplement for both years was ~86%. Self-response was required for the supplemental questionnaire (9,10).

Follow-up information on the NHIS respondents. The NCHS has collected information to allow linkage of the NHIS respondents ≥ 18 y old with the National Death Index (NDI) beginning with survey year 1986 (11). This linkage information allows matching of the NHIS respondents with NDI and other data collected by the NCHS. The linkage thus provides a longitudinal passive follow-up (respondents were not recontacted or reinterviewed) component to the NHIS that permits determination of vital status and allows estimation of survival, mortality, and life expectancy in relation to information collected in the core or supplemental NHIS questionnaires. Linkage information was available through Dec. 31, 1997.

The NCHS calibrated the NDI matching for NHIS against 2 studies in which vital status was ascertained by active follow-up (the National Health and Nutrition Examination Survey I Epidemiologic Follow-up Study and the Longitudinal Study of Aging). The NHIS matching methodology correctly classified 97% of true matches and 97% of false matches in these 2 studies (11).

Dietary assessment method. In both the NHIS 1987 and the NHIS 1992, a ~60-item semiquantitative FFQ, developed by Block and colleagues, was used to collect dietary information (12,13). The validation of this instrument included comparison of estimated intakes of energy and nutrients obtained from the FFQ with those obtained from multiple 24-h recalls and dietary records (12–14). The questionnaire included queries about frequency of consumption and portion size of food items consumed over the past year. The FFQ also included questions about dietary behaviors related to removal of poultry skin or fat from red meat. The FFQ used in 1992 contained a slightly larger number of food items than the one used in 1987. For the present study, only those items common to both survey years were considered for the combined analysis.

Analytic cohort. The potential analytic cohort for the present study included all respondents ≥45 y old, who completed the FFQ as part of the supplement to each survey (n = 14,769). From this potentially eligible sample, we excluded subjects who left >5 items of the FFQ incomplete or whose responses were considered invalid on the basis of previous validation studies (n = 1198) (12,13). Respondents with 1–5 incomplete FFQ items were considered to be nonconsumers of the respective items. The mean age and BMI of respondents excluded for incomplete or invalid FFQs were comparable to those retained in the cohort; however, more of those excluded provided no information on smoking status, level of education, and supplement use.

As recommended by the NCHS, respondents not eligible for an NDI match due to lack of information to allow matching were excluded from the analytic cohort (n = 229) (11). Finally, we excluded respondents who self-reported a history of chronic disease(s) (ischemic heart disease, cerebrovascular disease, any cancer except skin, diabetes, and hypertension) at baseline (n = 3171), or for whom the cause of death was accident or injuries (ICD-9 codes 800–999) (n = 87). With these exclusions, the analytic cohort comprised 4146 men and 5938 women for 1987 and 1992 combined. The final analytic cohort of 10,084 included 1005 deaths due to all causes in men and 1282 deaths in women, for a total 79,345 person-years.

Determination of dietary patterns

Dietary score-based patterns. Prevalent dietary guidance recommends the consumption of fruits, vegetables, whole grains, low-fat dairy, lean meat/poultry/fish and meat alternates such as beans/nuts/and seeds (3). Dietary guidance also recommends certain dietary behaviors to reduce saturated fat intake, e.g., removal of poultry skin and fat from red meats. For this study, we modified the previously reported RFS slightly to include meat alternates and behaviors related to consumption of fat from meat or poultry skins. Thus, we examined each FFQ for mention of food items that may be considered recommended foods and desirable dietary behaviors related to removal of poultry skin or fat from red meat. All decisions regarding these criteria for being considered a healthy food or behavior were made before beginning analysis. From the ~60-item FFQ, 23 items were found to correspond to our criteria of recommended foods and included the following: all fruits (orange juice or grapefruit juice; other fruit juices; oranges; grapefruits; cantaloupes; apples or applesauce), vegetables (carrots or mixed vegetables with carrots; tomatoes; green salad; broccoli; spinach; mustard/collard greens; baked or boiled potatoes; sweet potatoes), whole grains (cooked cereals such as oatmeal; high-fiber cereals; dark breads; corn tortillas and breads), low-fat or nonfat dairy (2 or 1% fat milk; skim milk), lean meats/poultry and alternates (baked or broiled chicken and turkey; dry beans; nuts), for a total of 23 items. The dietary behavior related questions in the FFQ included "When you eat chicken or other poultry, how often do you eat it with the skin on? Would you say often, sometimes, rarely, or never?" and, "When you eat red meat, how often do you eat the fat? Would you say often, sometimes, rarely, or never?" Mention of a recommended food at least 1 time/wk and mentioning that chicken skin or fat on red meat was never or rarely eaten contributed 1 point each to the overall score. We labeled the resulting dietary pattern a Recommended Foods and Behavior Score or RFBS, maximum = 25. The resulting score variable was operationalized as a continuous variable and as approximate weighted quartiles based on distribution across the entire cohort. The RFBS and RFS were highly correlated (r = 0.97).

Factor analysis–derived dietary patterns. The responses to the 23 recommended food items and 2 behavior-related questions were subjected to a principal components analysis (PCA), using PROC FACTOR available in the SAS package (14). We used the appropriate NCHS assigned sample weights for NHIS in the PCA procedures. We retained the first 3 components from this initial analysis on the basis of eigenvalues (>1.25), proportion of explained variance (≥5%), and interpretability. The eigenvalues for components 1–3 were, 3.1, 1.7, and 1.3, and these 3 factors explained 12.5, 6.9, and 5.5%, of the variance, respectively. The retained components were then rotated using varimax rotation to obtain an orthogonal solution. Fruit, vegetable, and whole-grain items had high factor loadings in the 1st factor and the factor was labeled as the fruit-vegetable-whole grain factor. The 2nd factor was labeled as the ethnic pattern (beans, corn bread/tortillas, and mustard greens loaded on this factor). The 3rd factor had high loadings for skim milk and behavior-related items and was labeled as the low-fat pattern. Factor scores were computed for the retained factors using the NFACT and OUT options in the FACTOR procedure (14). Briefly, this option computes each subject's factor score for the retained components by summing the product of the optimal regression weights with the subject responses to each item. The resulting score was then used as a continuous exposure variable and to form categories of exposure from the sample weighted quartiles of these scores.

Cluster analysis–derived dietary patterns. The responses to the 23 recommended foods items and 2 behavior-related questions were

TABLE 1

Sociodemographic characteristics of respondents by approximate weighted quartiles of the Recommended Foods and Behavior Score[1]

	Approximate quartiles of RFBS			
	Q1	Q2	Q3	Q4
	0–8	9–11	12–14	≥15
n	2219	2735	2866	2264
Age,[1] *y*	58.9 ± 0.3	60.0 ± 0.3	60.3 ± 0.3	61.1 ± 0.3
BMI,[1] *kg/m²*	25.9 ± 0.1	25.9 ± 0.1	25.7 ± 0.09	25.4 ± 0.1
RFBS[1]	6.3 ± 0.05	10.0 ± 0.02	12.9 ± 0.02	16.5 ± 0.04
	%			
Caucasian	82.3	85.0	86.0	85.8
Female	46.9	50.0	54.8	61.3
Aged ≥65 y	30.0	33.7	34.5	37.6
<12 y of education	40.4	30.2	23.8	20.5
Income < $50,000/y	68.9	65.9	63.3	61.6
Below poverty	9.3	6.7	4.6	4.4
Never smokers	32.4	39.9	46.5	51.3
Current smokers	36.9	27.0	20.7	14.9
Former smokers	30.7	33.1	32.8	33.8
BMI > 24.9	54.4	55.2	52.9	48.1
Reported supplement use	39.9	44.6	53.7	57.2
Drink alcohol	54.4	56.9	57.1	60.3

[1] Values are means ± SEM. For each variable in the table, the test of heterogeneity among RFBS quartiles was significant ($P < 0.0001$).

also used in a cluster analysis by applying the FASTCLUS procedure in SAS (15). This clustering procedure can efficiently determine a prespecified number of disjoint clusters for a large data set such as the NHIS. In this procedure, Euclidean distance is used to determine which observations are close enough to belong to the same cluster. All analyses were weighted with the appropriate NHIS sample weights. The procedure was used to form different cluster solutions ranging from 2 to 9 clusters. Considerable controversy exists regarding criteria used for determining the number of clusters (16). Most studies using cluster analysis choose the number of clusters after a review of results obtained from running a number of different solutions. We decided on a 4-cluster solution after examining the gain in explained variance, understandability of the clustered dietary characteristics, and cluster size for statistical power for subsequent outcome analysis.

The 4-cluster solution also allowed comparison with quartiles of RFBS and factor scores. The characteristics of the 4 clusters were as follows: cluster 1, less likely to mention whole grains, low-fat or skim milk, and to remove fat from meat and poultry; cluster 2, less likely to mention most fruits and vegetables; cluster 3, less likely to mention most fruits, and high-fiber cereals; cluster 4, highest proportion reporting weekly use of most items. We note that clusters obtained by this analysis do not represent an ordered variable. Therefore, trend tests with cluster-derived patterns are not valid. In our presentation of results in this study, we chose to order the clusters according to magnitude of risk reduction, with the least favorable serving as the reference for comparison.

Analytic methods

Computation of person-years. The date of entry was Wednesday of the week of interview within the quarter within the year of interview (1987 or 1992), and the exit date was 12/31/97 or date of death. When the date of death was given only as month and year, it was set as the 15th of the month of the given year.

Statistical analysis. The relative risk of mortality associated with categories of dietary patterns from various approaches was estimated using Cox's proportional hazards regression models with the SURVIVAL procedure available in the SUDAAN software package (17,18). SUDAAN software is designed for analysis of complex survey data (17). Age was used as the time metric for the survival analysis, and the baseline hazard in the proportional hazard regression was stratified by year of birth (19). The variables available for covariate adjustment included year of the survey (1987 vs. 1992), race/ethnicity, level of education (y), BMI computed from self-reported body weight and height, smoking status, supplement use, reported alcohol use, and total energy intake. The multivariate models to determine the independent association of dietary patterns with mortality included all of the variables mentioned above. The variables to be included in these models were decided a priori based on published information on established risk factors for poor health. All analyses were stratified by sex and were weighted for the sample weights. All statistical tests were two-tailed at the 0.05 level. Values presented in the text are means ± SE.

We also computed the population-attributable fraction using estimated multivariate- adjusted relative hazards from the Cox proportional hazard regression models and its sample design-based SE using the Taylor linearization method (Graubard, B. I. & Fears, T. R., unpublished data).

RESULTS

Sociodemographic characteristics of respondents. The mean age of the analytic cohort was 60.1 ± 0.16 y, and range was 45–90 y, at baseline. The median follow-up time was 5.9 y (mean, 7.26 ± 0.04). Respondents in the higher RFBS quartiles had a higher level of education and income, a higher proportion were never smokers, had a BMI < 25, and reported use of supplements ($P < 0.0001$) **(Table 1)**. The profiles of respondents reporting higher fruit-vegetable-whole grain factor scores or those in the 4th cluster were similar to those of respondents reporting a higher RFBS.[4] The RFBS and the fruit-vegetable-whole grain factor scores were positively related ($r = 0.84$). The mean RFBS was higher in the 4th relative to the other 3 clusters ($P < 0.0001$). RFBS and

[4] Appendices 1 and 2 present sociodemographic characteristics of respondents by quartiles of fruit-vegetable-whole grain factor scores, and by 4 clusters; they are available with the online posting of this paper at www.nutrition.org.

TABLE 2

Relative risk of all-cause mortality in relation to dietary patterns derived from 3 approaches, for men in the NHIS 1987, 1992[1]

Dietary pattern	Age-Survey year adjusted, $n = 4146$ (cases = 1005)	Multivariate adjusted,[2] $n = 4007$ (cases = 970)
Approximate weighted quartiles of Recommended Foods and Behavior Score		
0–8	1.0 (Reference)	1.0 (Reference)
9–11	0.87 (0.71–1.06)	0.88 (0.72–1.08)
12–14	0.77 (0.62–0.94)	0.84 (0.68–1.03)
≥15	0.64 (0.51–0.80)	0.72 (0.56–0.92)
P for trend	<0.0001	0.001
Weighted quartiles of factor 1 (fruit-vegetable-whole grain factor) score		
1st Quartile	1.0 (Reference)	1.0 (Reference)
2nd Quartile	0.85 (0.69–1.04)	0.92 (0.74–1.13)
3rd Quartile	0.74 (0.60–0.91)	0.84 (0.68–1.06)
4th Quartile	0.63 (0.50–0.79)	0.74 (0.57–0.95)
P for trend	<0.0001	0.002
Patterns from clusters		
Cluster 1	1.0 (Reference)	1.0 (Reference)
Cluster 2	0.82 (0.67–1.01)	0.94 (0.76–1.16)
Cluster 3	0.75 (0.61–0.93)[a]	0.87 (0.71–1.07)
Cluster 4	0.65 (0.54–0.79)[b]	0.82 (0.66–1.01)[c]

[1] Values are estimated relative risk (95% CI). Letters within a column indicate significant difference from the reference cluster, [a] $P = 0.009$; [b] $P < 0.0001$; [c] $P = 0.05$.

[2] The multivariate-adjusted models included: age, year of survey (1987 or 1992), race, BMI, education in years, smoking status, supplement use, alcohol use, and energy intake. Multivariate models include only those with complete covariate information.

fruit-vegetable-whole grain factor scores were related positively with reported energy, dietary fiber, and micronutrient intake but negatively with total and saturated fat intake ($P < 0.05$). Similarly, in the 4th cluster, the mean total and saturated fat intakes were lower but micronutrient intake was higher relative to other clusters ($P < 0.05$; data not shown).

Relative risk of all-cause mortality associated with various dietary patterns. In men, a strong inverse trend of risk of all-cause mortality ($P < 0.0001$) in relation to RFBS and fruit-vegetable-whole grain factor scores was noted in age-survey year-adjusted models (**Table 2**). Similarly, the risk was significantly lower in clusters 3 and 4 relative to the 1st cluster. After multivariate adjustment, the inverse association of RFBS and fruit-vegetable-whole grain factor scores with risk of mortality remained significant ($P \leq 0.002$), with the highest RFBS or fruit-vegetable-whole grain factor score quartiles showing risk reduction of ~26–28% relative to the lowest categories. In patterns derived from cluster analysis, the multivariate-adjusted risk was lower in the 4th cluster relative to the 1st cluster by 18% ($P = 0.05$).

In women, a strong inverse trend of risk of all-cause mortality ($P \leq 0.007$) in relation to RFBS and fruit-vegetable-whole grain factor scores was noted in age-survey year-adjusted models (**Table 3**). Similarly, the risk was significantly lower in

TABLE 3

Relative risk of all-cause mortality in relation to dietary patterns derived from 3 approaches, for women in the NHIS 1987, 1992[1]

Dietary pattern	Age-Survey year adjusted, $n = 5938$ (cases = 1282)	Multivariate adjusted,[2] $n = 5690$ (cases = 1217)
Approximate weighted quartiles of Recommended Foods and Behavior Score		
0–8	1.0 (Reference)	1.0 (Reference)
9–11	0.96 (0.78–1.19)	0.95 (0.76–1.18)
12–14	0.80 (0.64–0.98)	0.80 (0.64–1.02)
≥15	0.78 (0.63–0.97)	0.80 (0.61–1.04)
P for trend	0.005	0.04
Weighted quartiles of factor 1 (fruit-vegetable-whole grain factor) score		
1st Quartile	1.0 (Reference)	1.0 (Reference)
2nd Quartile	0.89 (0.74–1.08)	0.89 (0.72–1.08)
3rd Quartile	0.81 (0.66–1.01)	0.81 (0.64–1.02)
4th Quartile	0.81 (0.66–0.98)	0.87 (0.67–1.11)
P for trend	0.007	0.09
Patterns from clusters		
Cluster 1	1.0 (Reference)	1.0 (Reference)
Cluster 2	0.90 (0.73–1.10)	0.93 (0.75–1.16)
Cluster 3	0.90 (0.73–1.11)	0.93 (0.74–1.17)
Cluster 4	0.83 (0.69–1.00)[a]	0.88 (0.72–1.09)

[1] Values are estimated relative risk (95% CI). [a] Different from the reference cluster, $P = 0.04$.

[2] The multivariate-adjusted models included: age, year of survey (1987 or 1992), race, BMI, education in years, smoking status, supplement use, alcohol use, and energy intake. Multivariate models include only those with complete covariate information.

the 4th cluster relative to the 1st ($P = 0.04$). With multivariate adjustment, the risk estimates were attenuated and only the RFBS remained a significant predictor in these models. Women in the highest RFBS quartile had a 20% lower risk of mortality (P for trend 0.04). The multivariate-adjusted trends for the association of mortality with dietary patterns reflected in the ethnic factor (factor 2) or the low-fat factor (factor 3) were not significant (P for trend > 0.05) in men or women (data not shown). The population-attributable fraction for RFBS was 0.16 ± 0.08 in men and 0.09 ± 0.07 in women.

To exclude the possibility that the observed associations may be due to preclinical or unreported disease at baseline (which may affect dietary intake adversely and result in early mortality), we repeated the regression analyses mentioned above after exclusion of events occurring in y 1 of follow-up (193 deaths). The associations observed above were unchanged by these exclusions. We also examined the possibility that exclusion of respondents with incomplete or invalid FFQs ($n = 1198$) may have biased our results. In regression models that included respondents with FFQs determined to be invalid in our study, the association of dietary pattern variables and mortality remained significant ($P < 0.05$). Finally, we tested the interaction of dietary pattern variables and age by including an interaction term in regression models; in all cases, the interaction term was not significant.

Correspondence of healthy dietary patterns from 3 approaches. An approximately similar proportion of respondents mentioned the use of recommended foods and desirable dietary behaviors in the highest RFBS quartile, the highest quartile of fruit-vegetable-whole grain factor score, and the desirable dietary cluster (**Table 4**). Relative to the population average, respondents in the 4th quartiles of RFBS or the fruit-vegetable-whole grain factor scores or the 4th cluster had lower total fat and saturated fat intakes but higher micronutrient intakes (**Table 5**).

TABLE 4

Reported consumption of food items at least 1 time/wk by entire cohort and subgroups with the healthiest diet patterns, derived from 3 approaches, NHIS 1987, 1992[1]

	All	RFBS ≥ 15	4th Quartile factor 1[2]	4th Cluster
n	10084	2264	2521	3178
			%	
Orange/grapefruit juice	64.4	85.4	86.7	83.5
Other fruit juice	38.6	63.9	58.1	56.4
Oranges	37.3	66.6	70.0	67.4
Grapefruit	22.6	48.1	50.1	45.2
Cantaloupe (in season)	51.7	78.4	78.4	76.9
Apples	62.5	91.4	91.4	89.8
Carrots	66.1	91.6	94.1	89.7
Tomato	82.0	96.0	99.3	94.1
Salad	80.4	94.1	98.9	92.7
Broccoli	47.6	79.9	84.4	77.5
Spinach	20.2	42.1	40.8	37.3
Mustard/collard greens/kale	12.0	20.6	13.1	13.9
Baked potato	84.1	94.0	94.0	91.0
Yam/sweet potato	11.6	25.4	14.9	18.7
Baked/broiled chicken	49.9	75.8	73.0	74.6
Dry beans	42.2	60.5	50.4	53.6
Peanuts/peanut butter	40.7	56.5	46.2	47.1
Cooked cereal	31.2	56.2	39.1	47.7
High fiber cereal	41.0	69.3	60.7	65.6
Dark bread	62.4	86.3	85.3	83.1
Corn bread/tortillas	24.4	34.9	21.0	26.8
2% milk	26.5	34.5	32.2	31.6
Skim milk	20.5	36.9	29.6	29.5
Never/rarely eat fat on red meat	78.3	91.5	78.6	89.9
Never/rarely eat poultry skin	56.1	76.5	59.3	75.2

[1] Values are percentages.
[2] Fruit-vegetable-whole grain factor.

DISCUSSION

Our results support the notion that dietary patterns that included foods and behaviors recommended in current dietary guidance (fruits, vegetables, whole grains, low-fat dairy, lean meats and alternatives) were associated with a lower risk of mortality at 5.9 y of follow-up irrespective of the approach used for determining dietary patterns. Although the magnitude of risk reduction in association with dietary patterns was relatively modest, given the universality of dietary exposure, the population-attributable fraction is substantial. We estimate that ~16 and 9% of mortality from any cause in men and women, respectively, could be eliminated by adoption of desirable dietary behaviors reflected in the 4th quartile of RFBS.

To our knowledge, this is the first study to examine dietary patterns in relation to mortality in a nationally representative U.S. cohort while using appropriate sample weights and software to adjust for the differential probabilities of selection of subjects according to survey design and nonresponse. Previous studies were limited to specific nonrepresentative cohorts; for example, in the screening cohort of BCDDP (7), dietary patterns were examined in relation to all-cause mortality, and in The Nurses' Health Study and the Health Professionals' Study cohorts, the outcome was risk of major chronic diseases (20–22).

In this study, relative to the lowest RFBS quartile, the risk of mortality for men and women in the highest RFBS quartile was 28 and 20% lower, respectively. The magnitude of risk reduction in association with dietary patterns derived from the relatively simple scoring system used in the present study agrees with other such reports using different methods for deriving dietary patterns. Using a variety of indices to reflect dietary patterns, several studies reported an inverse association of diet quality indices and all-cause mortality (7,23–31). Two of these studies assessed the diet for Mediterranean diet characteristics in small cohorts of elderly men and women in Greece and Denmark and found a 17–20% reduction in smoking-adjusted risk of mortality with a 1 unit increase in the diet score (27,29). A recent study of >22,000 adults in Greece confirmed the risk reduction in those with patterns consistent with the traditional Mediterranean diet (26). A 13% reduction in age-, smoking-, and alcohol-adjusted risk of mortality in men ($n = 3045$) in the highest 3rd of a Healthy Diet Index was noted in selected cohorts of the Seven Countries Study (28). We found that women in the highest RFS quartile had a 30% lower risk of all-cause mortality (7). Michels and Wolk modified the RFS to suit the Swedish diet and reported a 42% lower risk of mortality in the highest quintile of women in a Swedish Mammography Screening cohort ($n = 59{,}038$) (31). Conversely, Osler et al. (32) ($n = 7316$) reported no association of their 4-point Healthy Food Index with all-cause mortality.

As is evident from Tables 4 and 5, the association of dietary patterns with risk reduction was stronger in men than women. Others have also reported similar sex/gender differences in the strength of the association of dietary patterns with health (20–23). The reasons for the observed differences in diet-associated risk of mortality may be biological or due to gender

TABLE 5

Nutrient intakes of entire cohort and subgroups with the healthiest diet patterns, derived from 3 approaches, NHIS 1987, 1992[1]

	All	RFBS ≥ 15	4th Quartile factor 1[2]	4th Cluster
n	10084	2264	2521	3178
Energy, *kJ*	5786 ± 67	6493 ± 63	6393 ± 63	6259 ± 50
Energy from saturated fat, %	12.8 ± 0.04	11.3 ± 0.1	11.7 ± 0.1	11.4 ± 0.1
Energy from total fat, %	37 ± 0.1	33 ± 0.2	34 ± 0.2	34 ± 0.2
Fiber, *g*	11 ± 0.1	15 ± 0.1	14 ± 0.1	14 ± 0.1
Vitamin C, *mg*	124 ± 1.1	181 ± 2.5	179 ± 2.3	174 ± 1.9
Vitamin E,[3] *α-TE*	7.4 ± 0.1	8.9 ± 0.1	8.6 ± 0.1	8.5 ± 0.1
Folate, *μg*	247 ± 1.4	316 ± 3.2	306 ± 3.0	303 ± 3
Carotenoids, *μg*	3087 ± 31	4651 ± 72	4340 ± 62	4302 ± 58
Calcium, *mg*	686 ± 5.2	820 ± 11	772 ± 9	765 ± 8

[1] Values are means ± SEM.
[2] Fruit-vegetable-whole grain factor.
[3] α-TE, α-tocopherol equivalents; RE, retinol equivalents.

differences in some unmeasured, uncontrolled confounder. These gender differences notwithstanding, the trend for a protective effect of patterns comprised of recommended foods and behaviors in women was in a similar direction to that of men.

Few studies have compared the relation of dietary patterns derived from various approaches with a specific outcome, after similar control for confounding in a cohort. The dietary patterns reflected in the RFBS are driven by existing diet-disease hypotheses, whereas the factor and cluster analyses in this study were data driven. To ensure comparability among methods, we limited the dietary variables used for factor and cluster analysis to those used for deriving the RFBS. Other investigators also make arbitrary decisions about collapsing FFQ items into a smaller number of groups before proceeding with factor analysis (5). In our study, the results from RFBS and factor analysis were comparable; however, the RFBS has the advantage of easy interpretability and reproducibility. Although results from other studies that compared dietary patterns from different approaches are not directly comparable to our results, it is still informative to examine these studies. Osler et al. (32) compared the predictive ability of dietary patterns derived from factor analysis with their 4-point Healthy Food Index. The pattern described as "prudent" in that study was a stronger predictor of mortality relative to the Healthy Food Index. In a recent evaluation of dietary patterns in association with hypertension in a large European cohort, factor analysis–derived patterns were unrelated to hypertension; however, the 3rd quartile of a score based on the DASH dietary patterns was a predictor of incident hypertension (33). McCann et al. (34) also found simple methods for categorizing foods into groups to be as successful as factors generated by PCA for discriminating between endometrial cancer cases and controls. Haveman-Nies et al. (35) found that a Mediterranean diet score was associated with dietary cluster–derived patterns in a predictable direction. In the present study, apart from the 4th cluster, the other 3 clusters showed little variation in RFBS.

The data from our study should be interpreted with the following limitations in mind. The NHIS data provide no information on the level of physical activity, an important confounder of the association of diet and disease. Inclusion of physical activity may attenuate the risk estimates reported herein. We note, however, that the multivariate regression models included both the BMI and total energy intake; together, the 2 may capture some of the variability in physical activity (or contribution of physical activity to mortality). Also, energy intake may at least partially reflect the level of physical activity (36). Nevertheless, due to the association of dietary behaviors with other lifestyle factors, the possibility of residual confounding due to poorly measured variables related to diet and mortality cannot be ruled out.

Finally, all retrospective methods of assessing diet including the FFQ used in this study are subject to measurement error, which may lead to misclassification of respondents into dietary exposure categories (37,38). However, in the present study, dietary patterns were determined from dietary variables operationalized as whether recommended foods were mentioned at least once weekly without consideration for frequency of use and portion size, components of FFQ especially likely to be misreported (39,40). The cognitive ability of respondents to provide the portion size and detailed frequency of use of FFQ items is believed to be questionable (39). Thompson et al. (41) reported that the percentage of consumers or nonconsumers of an FFQ item showed a high degree of agreement with consumers or nonconsumers identified from daily food reports; however, the monthly frequency of use of FFQ items showed systematic bias. A recent report confirms previous findings that portion size in the FFQ contributed a relatively limited amount of information to the percentage of explained variance in food intake (40,42). We have found the RFBS reported here and a version of the RFBS that incorporates weekly frequency of use of items to be highly correlated, with the RFBS computed from at least weekly mentions of food items explaining 71% of the variance in the RFBS computed from weekly frequency of FFQ items. Therefore, although our approach is unsuitable for quantitative estimates of food and nutrient intake, we speculate that it is somewhat successful in differentiating between people who consume or do not consume the recommended foods.

In conclusion, our data suggest that the recommendation for consuming fruits, vegetables, lean meat/poultry and alternatives, low-fat dairy, and whole grains has considerable merit for improving health outcomes in the U.S. population.

ACKNOWLEDGMENT

We thank Lisa Licitra Kahle, IMS, Silver Spring, MD, for expert computing assistance.

LITERATURE CITED

1. Committee on Diet and Health, National Research Council (1989) Diet and Health: Implications for Reducing Chronic Disease Risk. National Academy Press, Washington, DC.

2. US Department of Health and Human Services (1988) The Surgeon General's Report on Nutrition and Health. DHHS publication (PHS) 88–50210. U.S. Government Printing Office, Washington, DC.

3. United States Department of Agriculture (2000) Nutrition and Your Health: Dietary Guidelines for Americans, 5th ed. U.S. Home and Garden Bulletin no. 232, Washington, DC.

4. Kant, A. K. (1996) Indexes of overall diet quality: a review. J. Am. Diet. Assoc. 96: 785–791.

5. Kant, A. K. (2004) Dietary patterns and health outcomes. J. Am. Diet. Assoc. 104: 615–635.

6. Hoffman, K., Schulze, M. B., Boeing, H. & Altenberg, H. P. (2002) Dietary patterns: report of an international workshop. Public Health Nutr. 5: 89–90.

7. Kant, A. K., Schatzkin, A., Graubard, B. I. & Schairer, C. (2000) A prospective study of diet quality and mortality in women. J. Am. Med. Assoc. 283: 2109–2115.

8. National Center for Health Statistics (1999) Design and Estimation of the 1985–1994 National Health Interview Survey. Series 2, no. 110, Hyattsville, MD.

9. Schoenborn, C. A. & Marano, M. (1988) Current Estimates from the National Health Interview Survey, United States, 1987 [Vital Health Stat. 10 (166), DHHS publ. (PHS) 88–1594]. Department of Health and Human Services, Washington, DC.

10. Benson, V. & Marano, M. (1994) Current Estimates from the National Health Interview Survey, United States, 1992 [Vital Health Stat. 10 (189), DHHS publ (PHS) 94–1517]. Department of Health and Human Services, Washington, DC.

11. National Center for Health Statistics (1997) Data File Documentation. National Health Interview Survey. Multiple Cause of Death, Public Use Data File. 1986–1994 Survey Years. Hyattsville, MD.

12. Block, G., Hartman, A. H. & Naughton, D. (1990) A reduced dietary questionnaire: development and validation. Epidemiol. 1: 58–64.

13. Smucker, R., Block, G., Coyle, L., Harvin, A. & Kessler, L. (1989) A dietary and risk factor questionnaire and analysis system for personal computers. Am. J. Epidemiol. 129: 445–449.

14. Block, G., Thompson, F. E., Hartman, A. M., Larkin, F. A. & Guire, K. E. (1992) Comparison of two dietary questionnaires validated against multiple dietary records collected during a 1-year period. J. Am. Diet. Assoc. 92: 686–693.

15. SAS Institute Inc. (2000) SAS User's Guide. Release 8.1. SAS Institute, Cary, NC.

16. Tibshirani, R., Walther, G. & Hastie, T. (2001) Estimating the number of clusters in a data set via the gap statistic. J. R. Stat. Soc. B 63: 411–423, Part 2.

17. Shah, B. V., Barnwell, B. G. & Bieler, G. S. (1997) SUDAAN User's manual. Release 7.5. Research Triangle Institute, Research Triangle Park, NC.

18. Cox, D. R. & Oakes, D. (1984) Analysis of survival data. Chapman and Hall, London, UK.

19. Korn, E. L. & Graubard, B. I. (1999) Analysis of Health Surveys, pp. 117–118. John Wiley & Sons, New York, NY.

20. McCullough, M. L., Feskanich, D., Stampfer, M. J., Rosner, B. E., Giovanucci, E. L., Rimm, E. B., Hu, F. B., Hunter, D. J., Spigelman, D., Colditz, G. A. & Willett, W. C. (2002) Diet quality and major chronic disease risk in men and women: moving toward improved dietary guidance. Am. J. Clin. Nutr. 76: 1261–1271.

21. McCullough, M. L., Feskanich, D., Stampfer, M. J., Rosner, B. E., Hu, F. B., Hunter, D. J., Variyam, J. N., Colditz, G. A. & Willett, W. C. (2000) Adherence to the Dietary Guidelines for Americans and risk of major chronic diseases in women. Am. J. Clin. Nutr. 72: 1214–1222.

22. McCullough, M. L., Feskanich, D., Rimm, E. B., Giovannucci, E. L., Ascherio, A., Variyam, J. N., Spiegelman, D., Stampfer, M. J. & Willett, W. C. (2000) Adherence to the Dietary Guidelines for Americans and risk of major chronic diseases in men. Am. J. Clin. Nutr. 72: 1223–1231.

23. Kant, A. K., Schatzkin, A., Harris, T. B., Ziegler, R. & Block, G. (1993) Dietary diversity and subsequent mortality in the First National Health and Nutrition Examination Survey Epidemiologic Follow-Up Study. Am. J. Clin. Nutr. 57: 434–440.

24. Nube, M., Kok, F. J., Vandenbroucke, J. P., van der Heide-Wessel, C. & van der Heide, R. M. (1987) Scoring of prudent dietary habits and its relation to 25-year survival. J. Am. Diet. Assoc. 87: 171–175.

25. Murphy, S. P., Davis, M. A., Neuhaus, J. M. & Lein, D. (1996) Dietary quality and survival among middle-aged and older adults in the NHANES I Epidemiologic follow-up study. Nutr. Res. 16: 1641–1650.

26. Trichopoulou, A., Costacou, T., Bamia, C. & Trichopoulos, D. (2003) Adherence to a Mediterranean diet and survival in a Greek population. N. Engl. J. Med. 348: 2599–2608.

27. Trichopoulos, A., Kouris-Blazos, A., Wahlquist, M. L., Gnardelis, C., Lagiou, P., Polychronopoulos, E., Vassilakou, T., Lipworth, L. & Trichopoulos, D. (1995) Diet and overall survival in elderly people. Br. Med. J. 311: 1457–1460.

28. Huijbregts, P., Feskens, E., Raasnen, L., Fidanza, F., Nissinen, A., Menotti, A. & Kromhout, D. (1997) Dietary pattern and 20-y mortality in elderly men in Finland, Italy, and the Netherlands: longitudinal cohort study. Br. Med. J. 315: 13–17.

29. Osler, M. & Schroll, M. (1997) Diet and mortality in a cohort of elderly people in North European community. Int. J. Epidemiol. 26: 155–159.

30. Kumagai, S., Shibata, H., Watanabe, S., Suzuki, T. & Haga, H. (1999) Effect of food intake on all-cause mortality in the community elderly: a 7-year longitudinal study. J. Nut. Health Aging 3: 29–33.

31. Michels, K. B. & Wolk, A. (2002) A prospective study of variety of healthy foods and mortality in women. Int. J. Epidemiol. 31: 847–854.

32. Osler, M., Heitmann, B. L., Gerdes, L. U., Jorgensen, L. M. & Schroll, M. (2001) Dietary patterns and mortality in Danish men and women: a prospective observational study. Br. J. Nutr. 85: 219–225.

33. Schulze, M. B., Hoffmann, K., Kroke, A. & Boeing, H. (2003) Risk of hypertension among women in the EPIC Potsdam Study: comparison of relative risk estimates for exploratory and hypothesis-oriented dietary patterns. Am. J. Epidemiol. 158: 365–373.

34. McCann, S. E., Weiner, J., Graham, S. & Freudenheim, J. L. (2001) Is principal components analysis necessary to characterize dietary behavior in studies of diet and disease? Public Health Nutr. 4: 903–908.

35. Havemann-Nies, A., Tucker, K. L., de Groot, L.C.P.G.M., Wilson, P.W.F. & van Staveren, W. A. (2001) Evaluation of dietary quality in relationship to nutritional and lifestyle factors in elderly people of the US Framingham Heart Study and the European SENECA study. Eur. J. Clin. Nutr. 55: 870–880.

36. Willett, W. & Stampfer, M. (1998) Implications of total energy intake for epidemiologic analysis. In: Nutritional Epidemiology (Willett W., ed.), pp. 273–301. Oxford University Press, New York, NY.

37. Bingham, S. A. (1987) The dietary assessment of individuals, new techniques and recommendations. Nutr. Abstr. Rev. 57: 704–742.

38. Sempos, C. T. (1992) Some limitations of semiquantitative food frequency questionnaires. Am. J. Epidemiol. 135: 1127–1132.

39. Smith, A. F. (1993) Cognitive psychological issues of relevance to the validity of dietary reports. Eur. J. Clin. Nutr. 47: S6–S18.

40. Flegal, K. M., Larkin, F. A., Metzner, H. L., Thompson, F. E. & Guire, K. E. (1988) Counting calories: partitioning energy intakes from a food frequency questionnaire. Am. J. Epidemiol. 128: 749–760.

41. Thompson, F. E., Subar, A. F., Brown, C. C., Smith, A. F., Sharbaugh, C. O., Jobe, J. B., Mittl, B., Gibson, J. T. & Ziegler, R. G. (2002) Cognitive research enhances accuracy of food frequency questionnaire reports: results of an experimental validation study. J. Am. Diet. Assoc. 102: 212–225.

42. Noethlings, U., Hoffman, K., Bergmann, M. M. & Boeing, H. (2003) Portion size adds limited information on variance in food intake of participants in the EPIC-POTSDAM study. J. Nutr. 133: 510–515.